Identifier les scénarios complexes de colonisation à l'échelle continentale des champignons pathogènes à l'aide des méthodes ABC

Benoit Barrès[1,2,3], Stéphanie Robert[2], Catherine Abadie[2], Jean Carlier[2], Cyril Dutech[1] et Virginie Ravigné[2]

[1] UMR BIOGECO [2] UMR BGPI [3] UPR 31

Why reconstruct routes of emergences ?

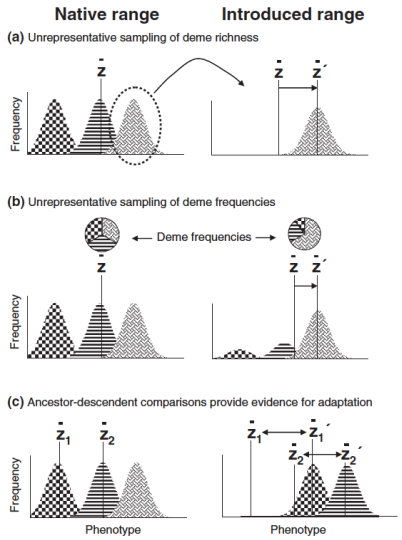
-Identify source populations

-Gain insights on dispersal mode

-Understand the process involved in emergence (admixture, bridgehead effect)

➔Improve quarantine politics

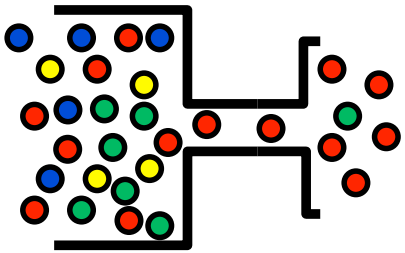
➔First step in the identification of evolutionary changes linked with emergence (comparison between source populations and emerging populations)



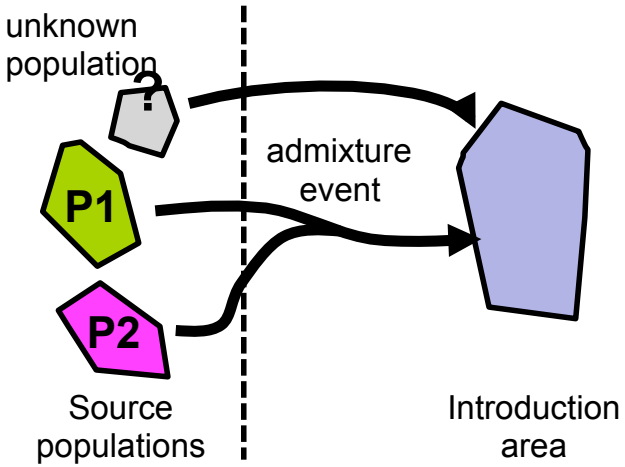
(Keller and Taylor, 2008)

Why use ABC methods ?

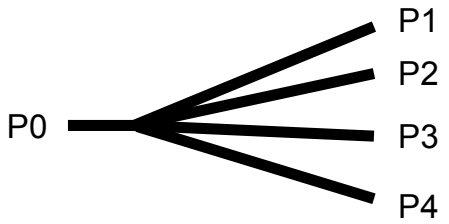
Demographic variation



Source population identification



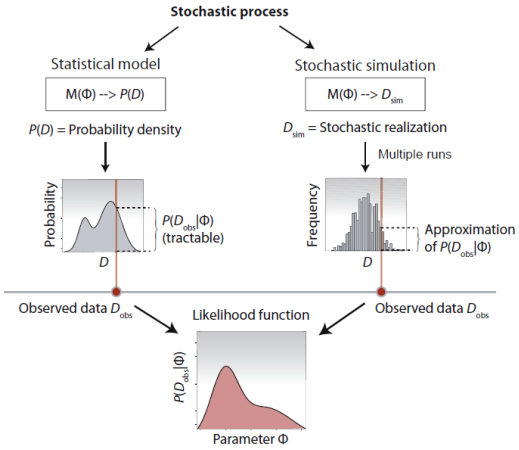
Multifurcation



-Assignment methodology (GeneClass2)
-Model-based coalescent method (MIGRATE, IM, …)

➔Need tractable Likelihoods

…but Likelihoods are not always tractable



(Hartig *et al.*, 2011)

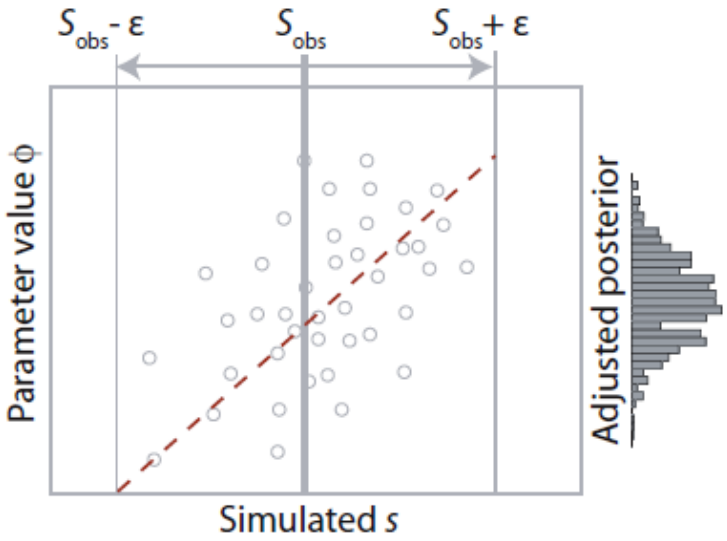
The ABC framework

3 steps :

1) generating (a lot of) simulated data sets

2) selecting simulated data sets closest to observed data set

3) estimating posterior distributions of parameters through a regression procedure



(*Excoffier et al.*, 2005)

(Hartig *et al.*, 2011)

Reviews: Beaumont 2010; Bertorelle *et al.*, 2010; Csilléry *et al.*, 2010

DIYABC software

-design and simulation of scenarios
(admixture, unsampled population, divergence between populations)

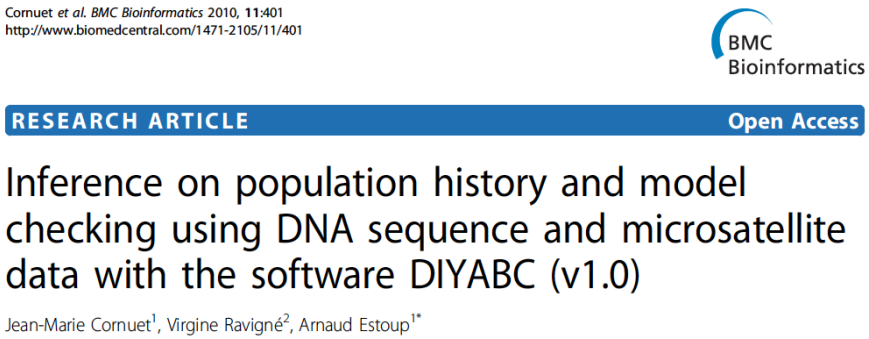
-tools to perform tests on priors distribution

-evaluation of posterior distribution

-tools to assess confidence and goodness of fit of the model

-haploid and diploid data

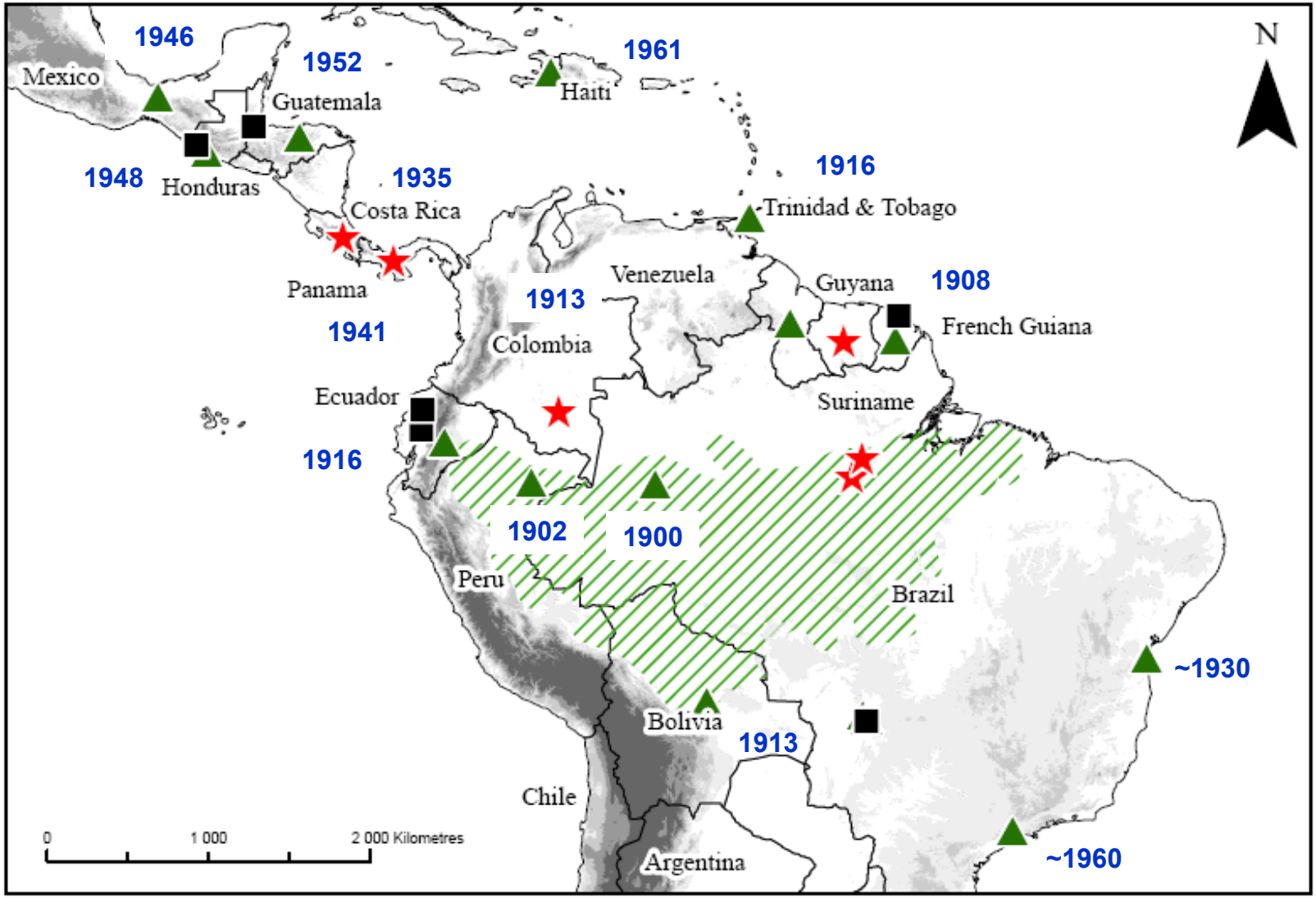
-microsatellite and/or sequence data



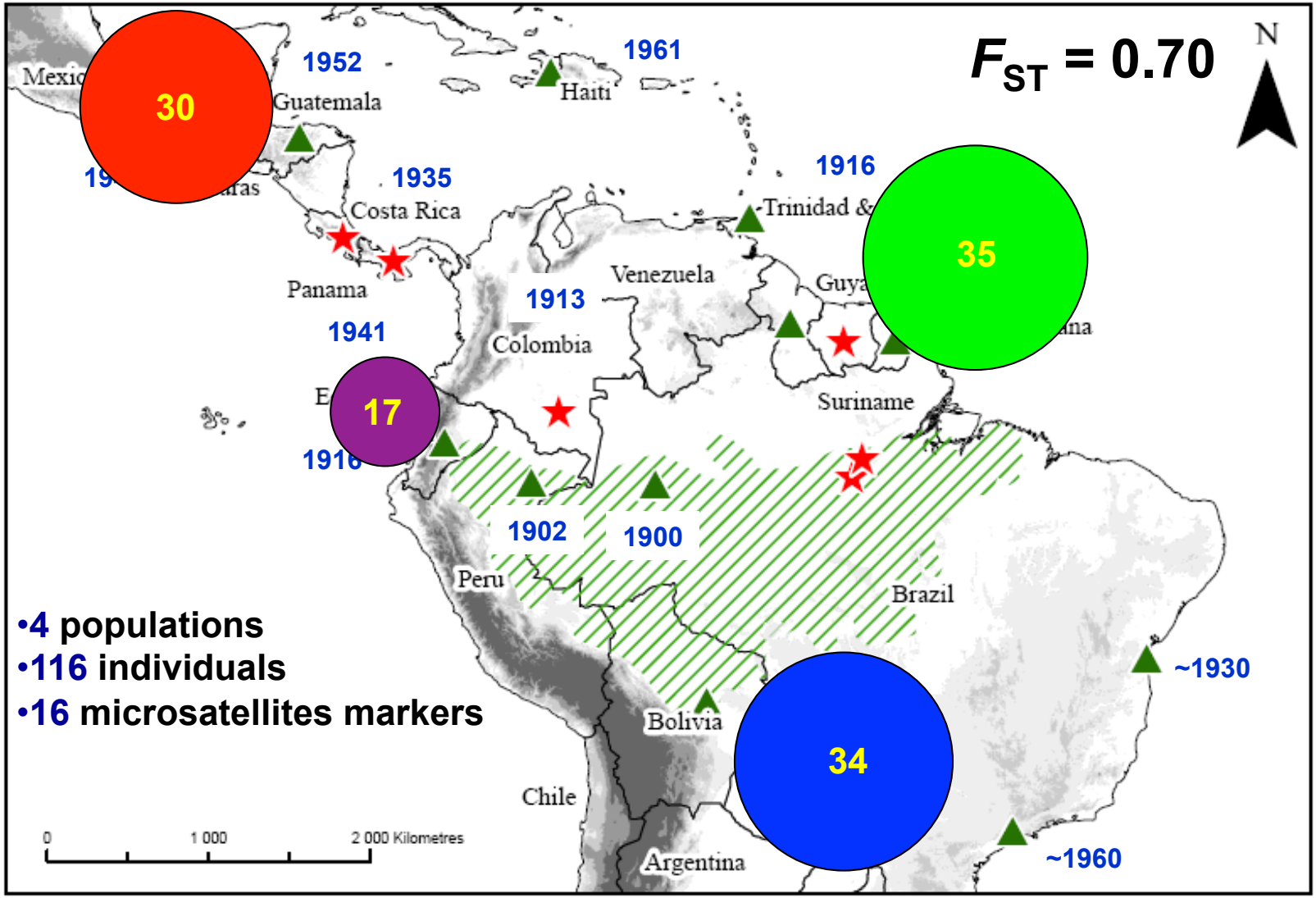
Continental scale dispersal of *Microcyclus ulei*

Benoit Barrès, Jean Carlier, Marc Seguin, Catherine Fenouillet, Christian Cilas and Virginie Ravigné

Continental scale dispersal of *Microcyclus ulei*

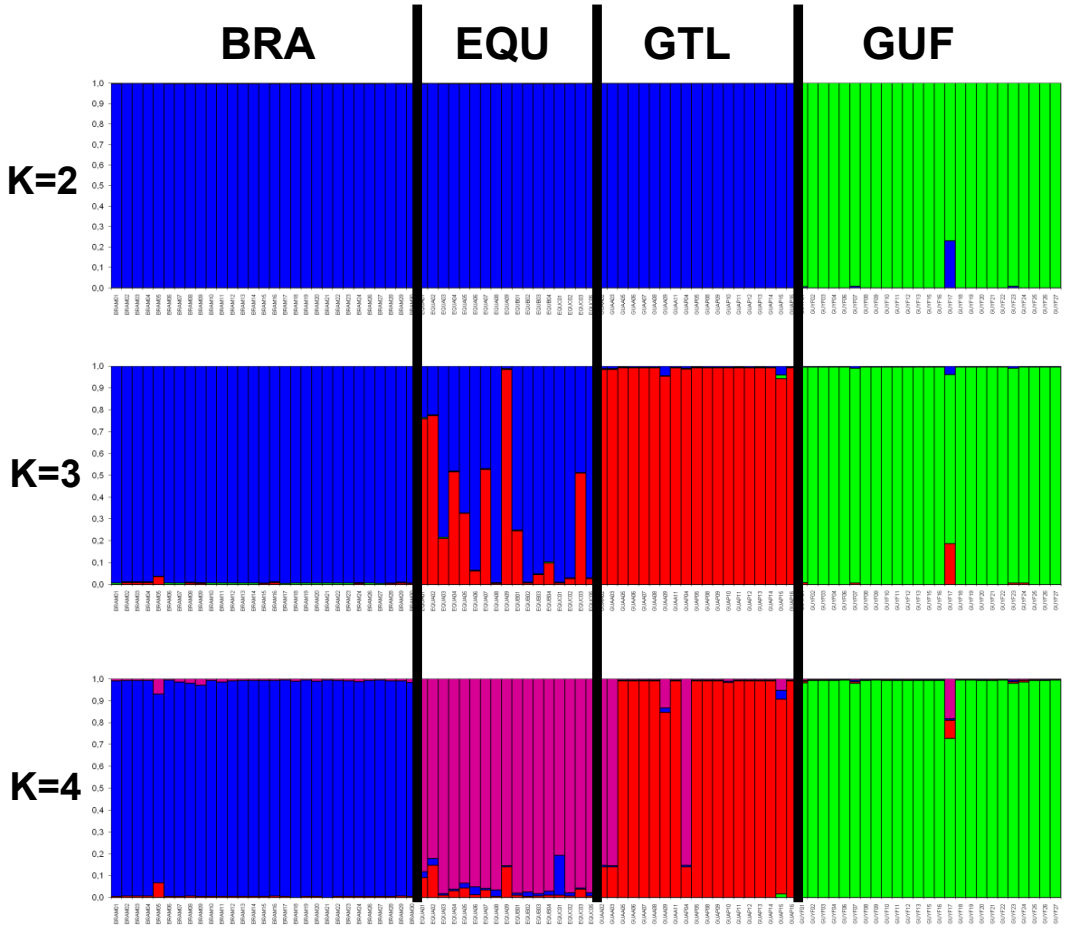


Sampling design

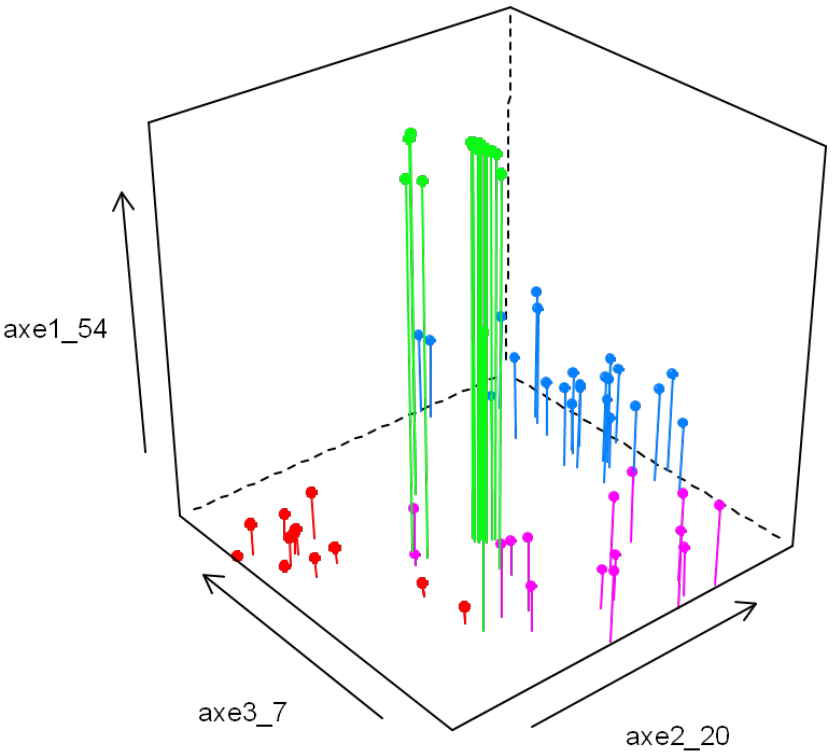


Genetic structure

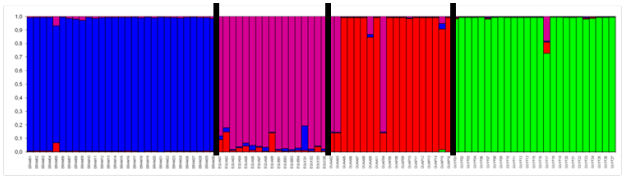
Bayesian Clustering

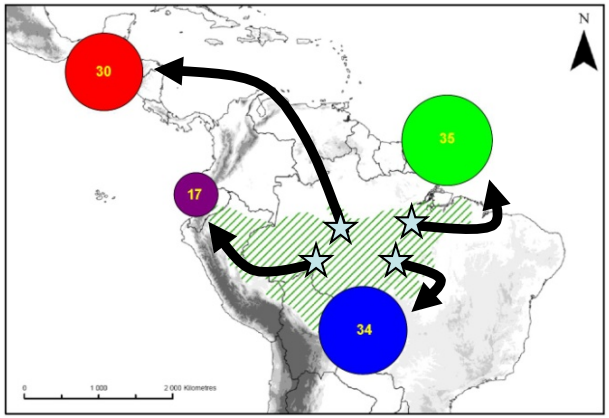


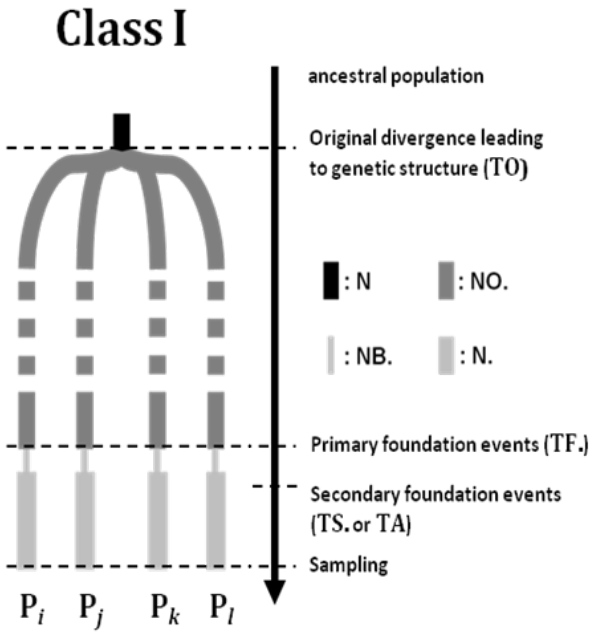
PCA



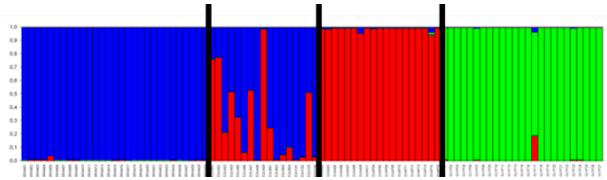
Four independent foundation events

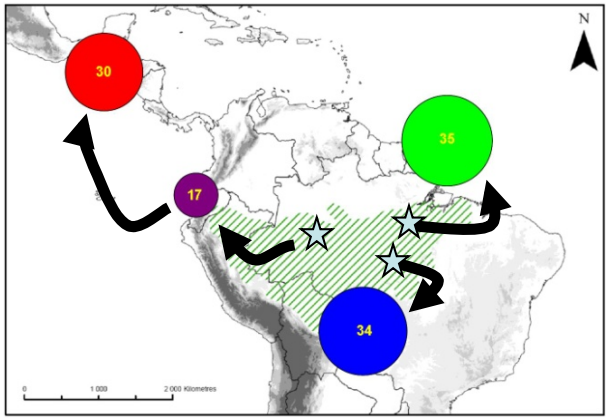


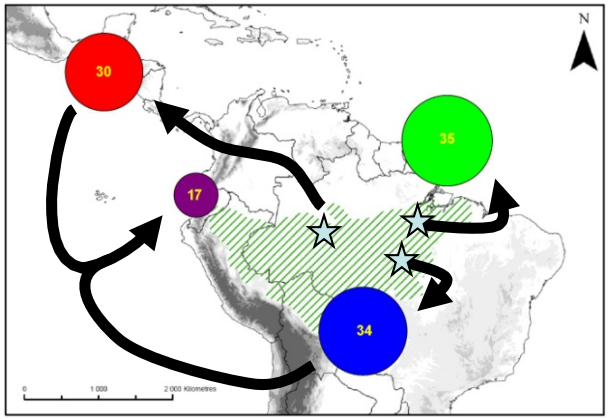


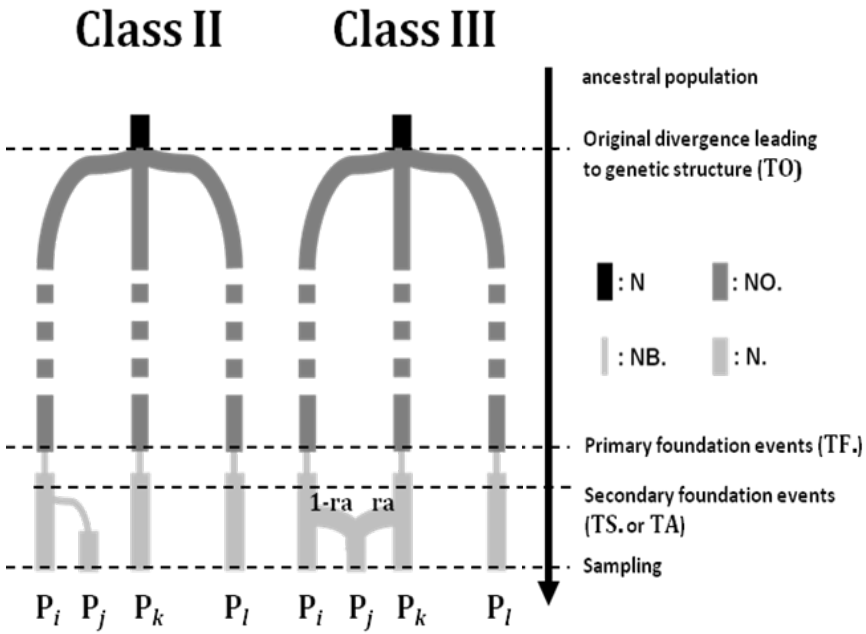


Three independent foundation events

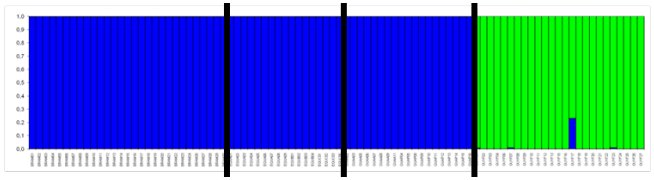


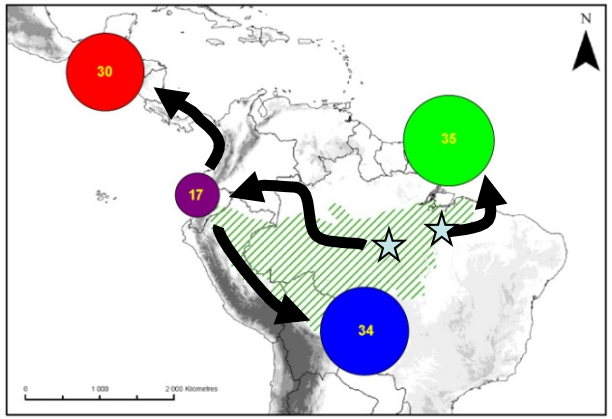


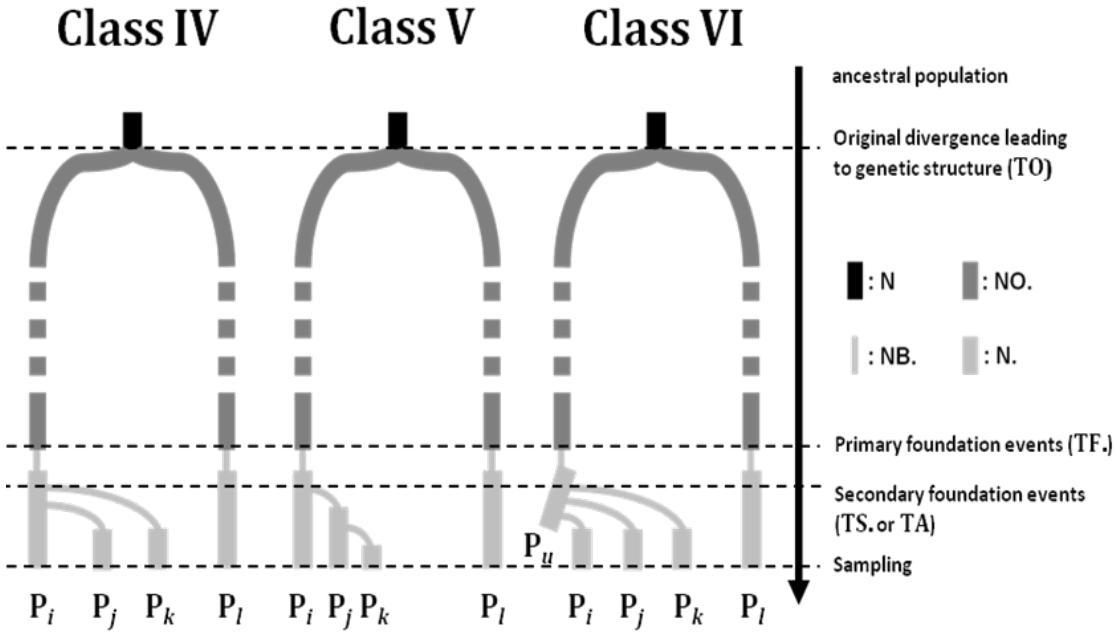


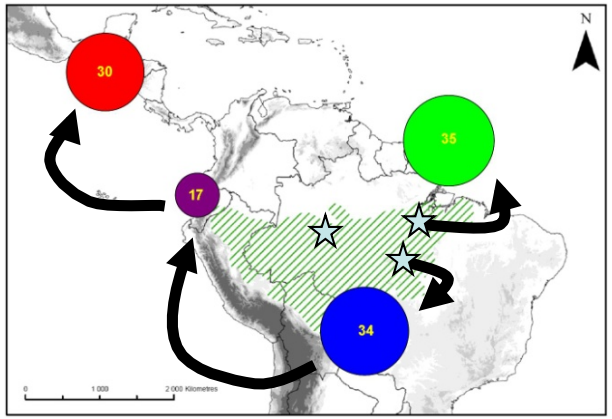


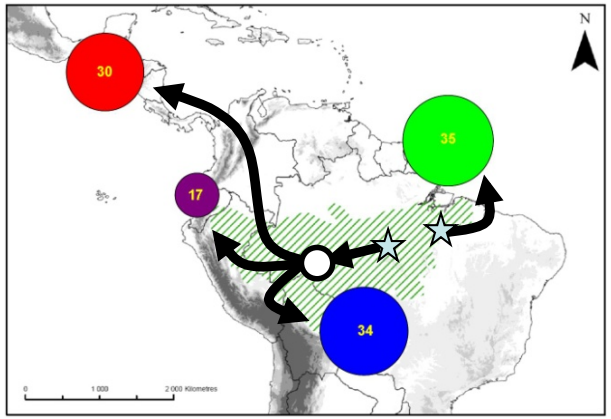
Two independent foundation events



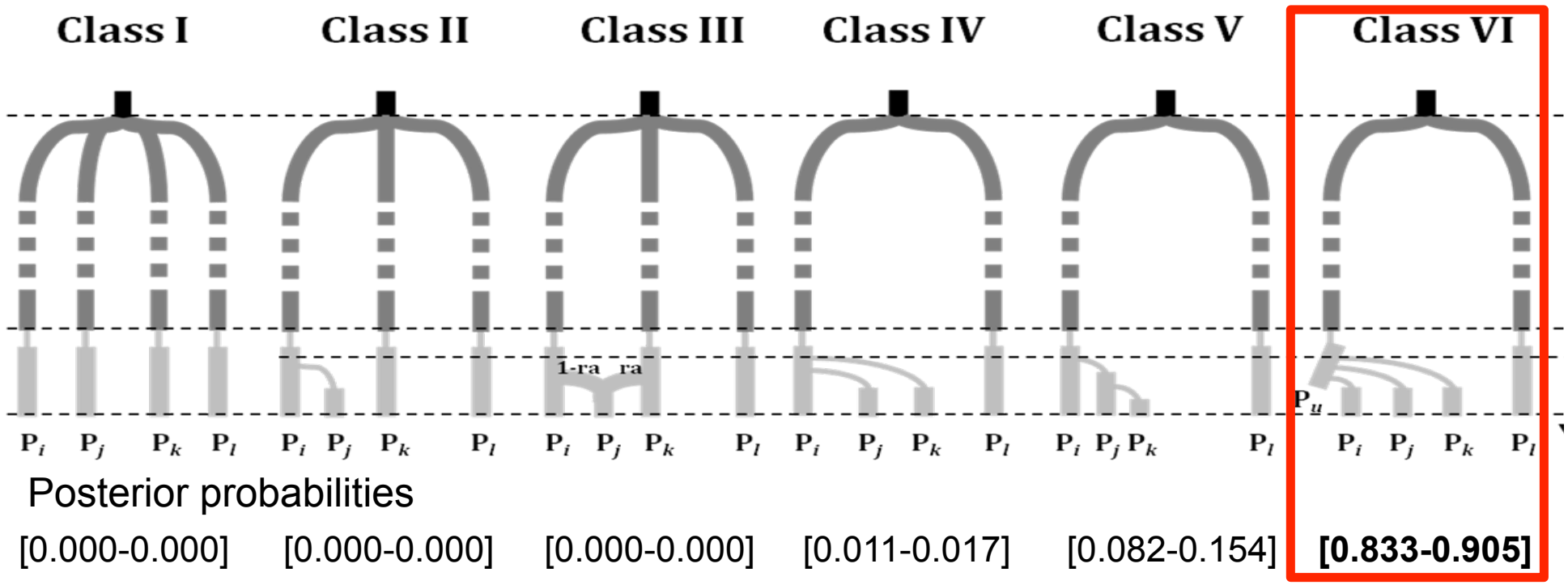








Continental scale dispersal of *Microcyclus ulei*



- Two independent foundation events
- An unknown common origin in western *Hevea* plantations

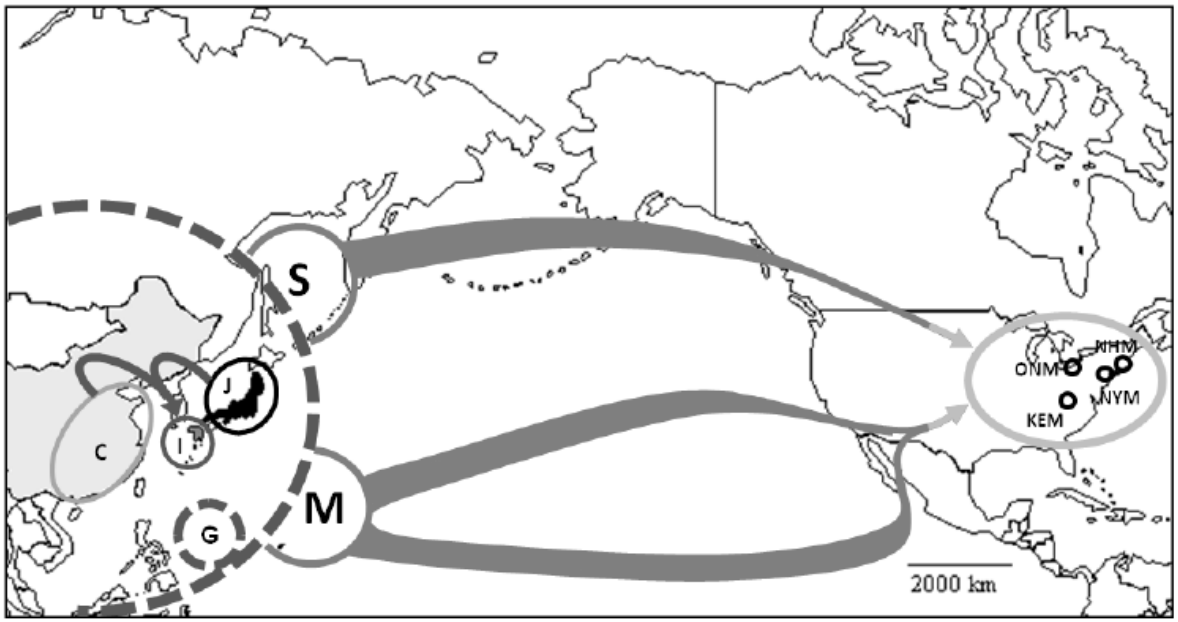
➔ man-mediated dispersal

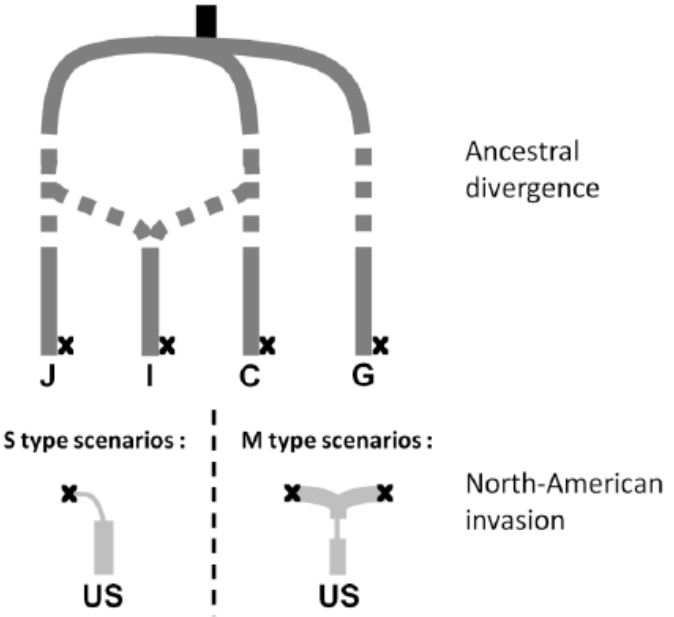
(Barrès *et al.*, submitted)

Introduction of *Cryphonectria parasitica* in North America

Cyril Dutech, Benoit Barrès, Julien Bridier, Cécile Robin, Michael Milgroom and Virginie Ravigné

Introduction of *Cryphonectria parasitica* in North America





-several hypotheses on the introduction in North America exist

-10 scenarios tested (4 with simple and 6 with multiple introduction)

Introduction of *Cryphonectria parasitica* in North America

Posterior probabilities of best tested scenarios :

	Japan	Japan + China	Japan + Intermediate	Japan + Unknown
NYM	**0.34** [0.31-0.38]	0.24 [0.22-0.27]	0.17 [0.15-0.19]	0.21 [0.19-0.24]
KEM	0.03 [0.03-0.04]	0.08 [0.07-0.10]	0.19 [0.17-0.22]	**0.51** [0.47-0.54]
NHM	0.14 [0.12-0.16]	0.14 [0.12-0.16]	0.17 [0.14-0.19]	**0.47** [0.43-0.50]
ONM	0.16 [0.14-0.19]	0.07 [0.06-0.08]	0.22 [0.20-0.25]	**0.39** [0.35-0.42]

- Corroborates the Japanese origin of *C. parasitica*
- No (or undetectable) Chinese introduction
- A possible admixture with an unidentified population

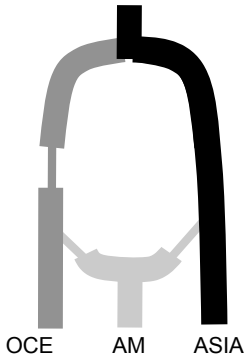
(Dutech *et al.*, accepted)

Introduction of *Mycosphaerella fijiensis* in South America

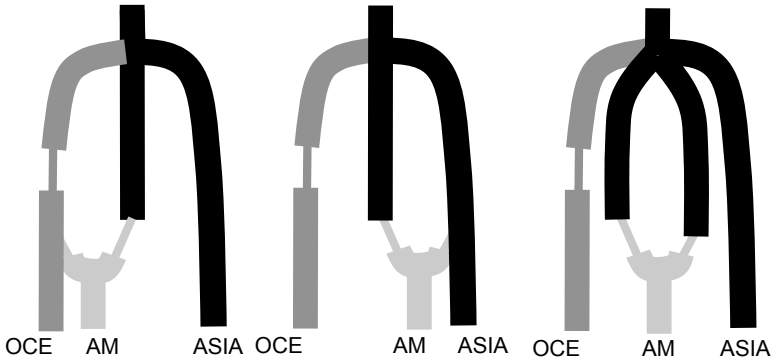
Stéphanie Robert, Jean Carlier, Marie-Françoise Zapater, Catherine Abadie and Virginie Ravigné

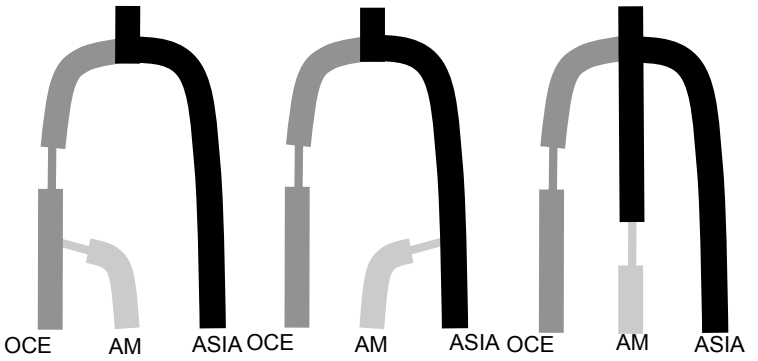
Introduction of *Mycosphaerella fijiensis* in South America

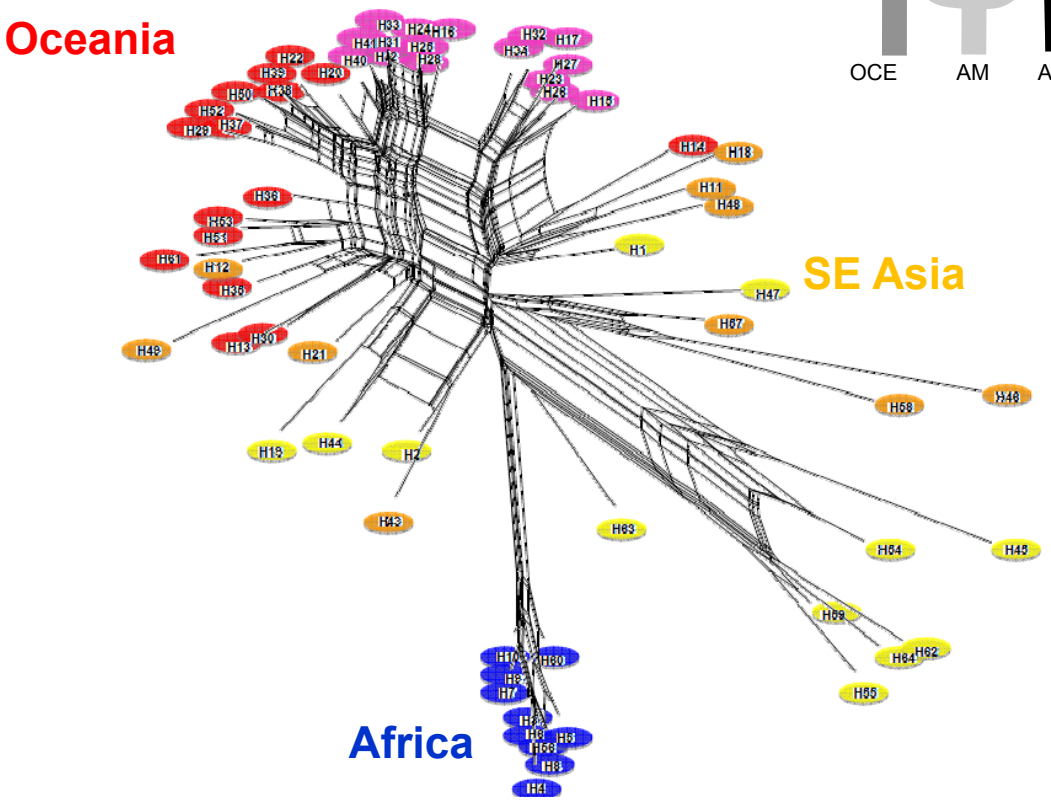
Scenario assumed:



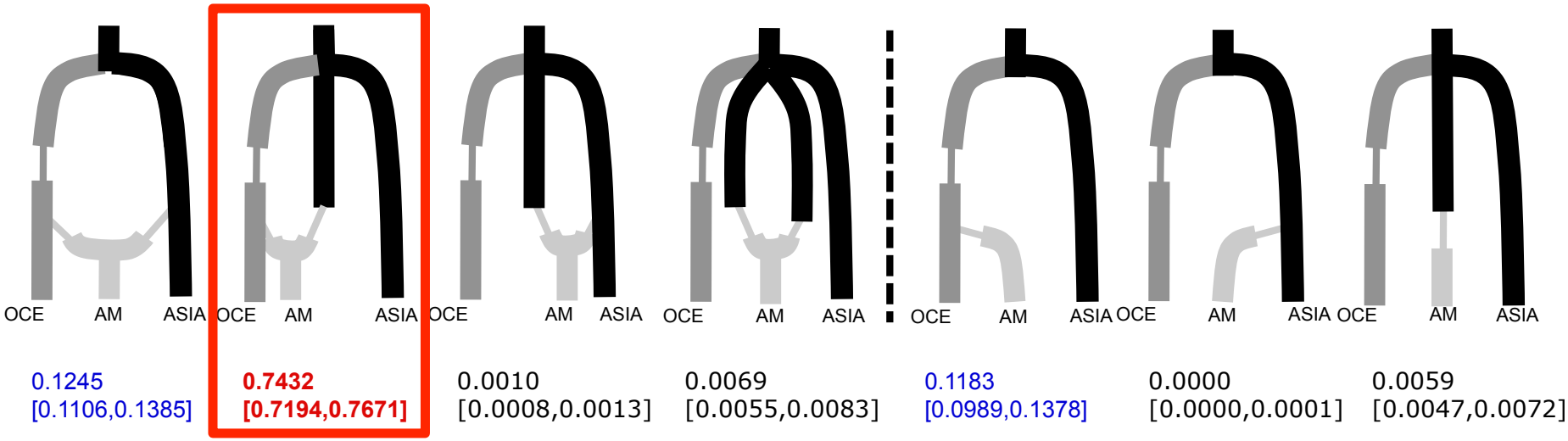
Other possible scenarios:







An introduction with admixture



- Introduction with admixture
- One unidentified source population

Conclusion

Method

ABC is a useful tool to decipher complex colonization scenario

-identification of admixture event

-importance of unsampled population

Biology

-identify source population (confirm or infirm historical knowledge)

-multiple introductions

-underline lack of knowledge on native area of pathogen

Some limitations

-panmictic population
-no regular gene flow between populations
-comparison between tested scenarios only

Thank you for your attention

UMR BIOGECO

Cyril Dutech
Cécile Robin
Julien Bridier

UMR BGPI

Jean Carlier
Virginie Ravigné
Elisabeth Fournier
Didier Tharreau
Daniel Bieysse

Cornell University

Michael Milgroom

UPR 31

Catherine Fenouillet
Christian Cilas
Jean Guyot
Franck Rivano

MICHELIN

Eric Cavaloc
Carlos Mattos
Milton Luiz

UMR DAP

Marc Seguin
Vincent Le Guen

Funding : CIRAD, MICHELIN, ANR EMERFUNDIS, ANR EMILE

Prior distribution of parameters

	Parameter name	distribution	Extremum values
Population effective sizes of the			
ancestral population	N	uniform	{10 – 100,000}
population leading to Brazilian population foundation	NOb	uniform	{10 – 100,000}
population leading to Equatorian population foundation	NOe	uniform	{10 – 100,000}
population leading to Guatemalan population foundation	NOgt	uniform	{10 – 100,000}
population leading to French Guyanian population foundation	NOgy	uniform	{10 – 100,000}
population leading to unsampled population foundation	NOu	uniform	{10 – 100,000}
bottleneck during foundation of Brazilian population	NBb	log-uniform	{1 - 100}
bottleneck during foundation of Equatorian population	NBe	log-uniform	{1 - 100}
bottleneck during foundation of Guatemalan population	NBgt	log-uniform	{1 - 100}
bottleneck during foundation of French Guyana population	NBgy	log-uniform	{1 - 100}
bottleneck during foundation of unsampled population	NBu	log-uniform	{1 - 100}
Brazilian population	Nb	uniform	{10 – 100,000}
Equatorian population	Ne	uniform	{10 – 100,000}
Guatemalan population	Ngt	uniform	{10 – 100,000}
French Guyana population	Ngy	uniform	{10 – 100,000}
Unsampled population	Nu	uniform	{10 – 100,000}
Time of the original divergence leading to genetic structure in source population	TO	uniform	{600 – 100,000}
foundation of the Brazilian population	TFb	log-uniform	{16 - 500}
foundation of the Equatorian population	TFe	log-uniform	{16 - 500}
foundation of the Guatemalan population	TFgt	log-uniform	{16 - 500}
foundation of the French Guyana population	TFgy	log-uniform	{16 - 500}
foundation of the unsampled population	TFu	log-uniform	{16 - 500}
secondary foudation of Brazilian population	TSb	log-uniform	{16 - 500}
secondary foudation of Equatorian population	TSe	log-uniform	{16 - 500}
secondary foudation of Guatemalan population	TSgt	log-uniform	{16 - 500}
secondary foudation of French Guyana population	TSgy	log-uniform	{16 - 500}
admixture event	TA	log-uniform	{16 - 500}

Population effective size :

-population UNIF {10-100,000}
-bottleneck LOG-U {1-100}

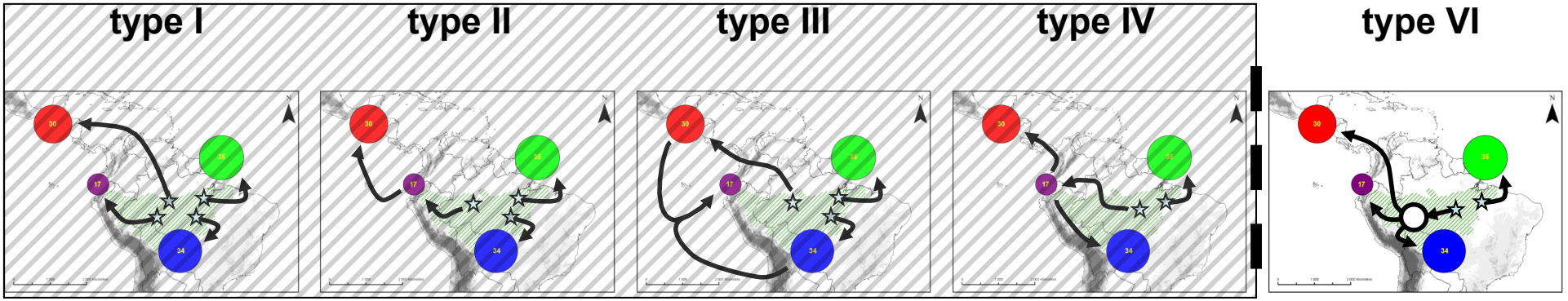
Time of divergence :

-ancestral UNIF {600-100,000}
-foundation pop LOG-U {16-500}

Admixture rate :

- UNIF {0.001-0.999}

The choice of a scenario and its consistency

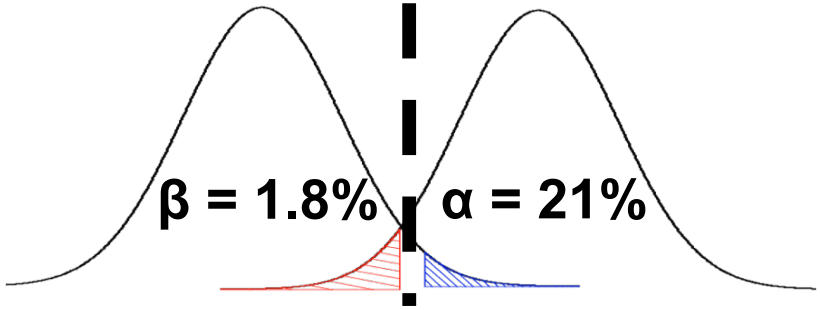


***P*=0.123**
[0.082-0.154]

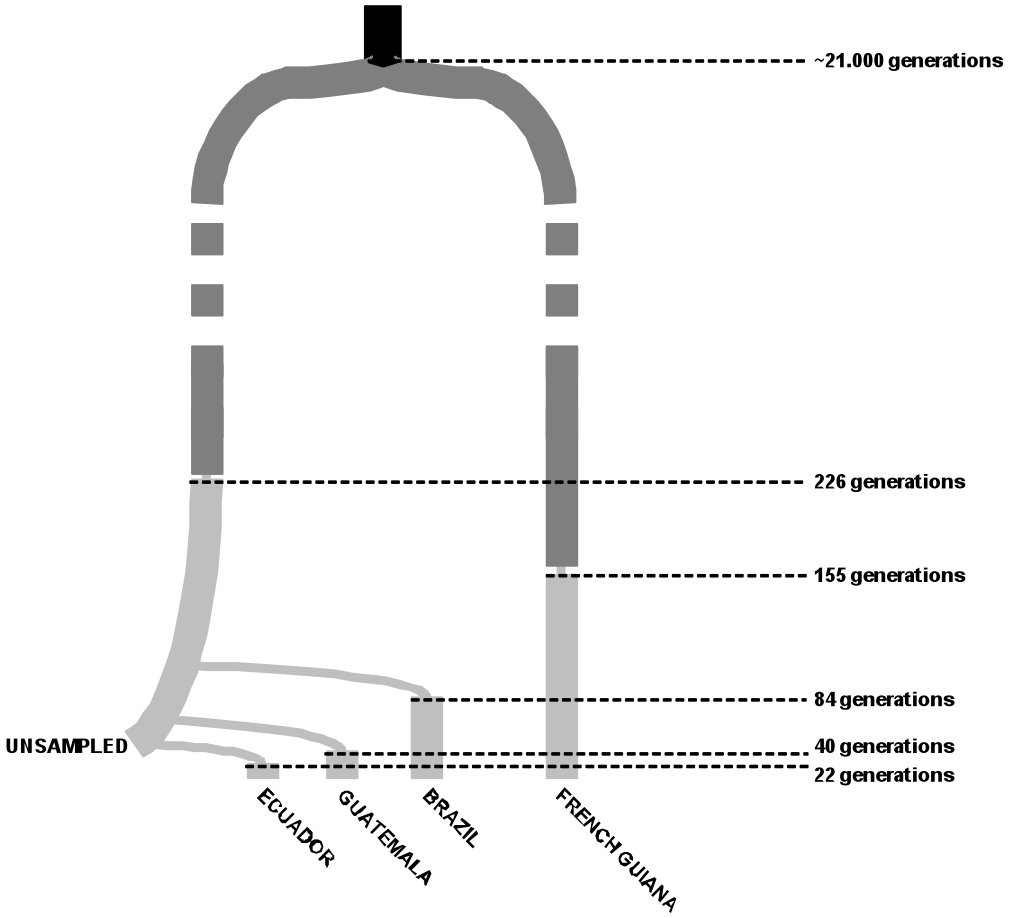
***P*=0.877**
[0.833-0.905]

-Goodness of fit evaluated by Model-checking

-Confidence in the choice of scenario evaluated with pseudo-observed datasets



Final scenario for *M. ulei*

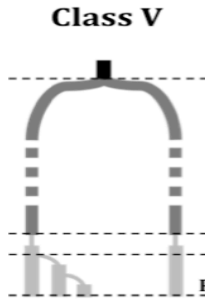


Erreur de Type II = proportion de jeux de données simulés sous de mauvais scénarios et qui pointent vers le meilleur scénario

Ici : 1.8 % !!

Erreur de Type I = proportion de jeux de données simulés sous le bon scénario et qui pointent vers un autre scénario

Ici : 21 % (beurk !)



Toutes les erreurs sont entre ce scénario et le scénario V

Final scenario for *M. fijiensis*